Identification and quantification of major faba bean seed proteins

Warsame, A. O. ORCID: https://orcid.org/0000-0002-9281-0443, Michael, N., O'Sullivan, D. M. ORCID: https://orcid.org/0000-0003-4889-056X and Tosi, P. (2020) Identification and quantification of major faba bean seed proteins. Journal of Agricultural and Food Chemistry, 68 (32). pp. 8535-8544. ISSN 1520-5118 doi: https://doi.org/10.1021/acs.jafc.0c02927 Available at https://centaur.reading.ac.uk/92133/

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

pubs.acs.org/JAFC

Article

Identification and Quantification of Major Faba Bean Seed Proteins

Ahmed O. Warsame, Nicholas Michael, Donal M. O'Sullivan, and Paola Tosi*

Cite This: https://dx.doi.org/10.1021/acs.jafc.0c02927

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Faba bean (*Vicia faba L.*) holds great importance for human and animal nutrition for its high protein content. However, better understanding of its seed protein composition is required in order to develop cultivars that meet market demands for plant proteins with specific quality attributes. In this study, we screened 35 diverse *Vicia faba* genotypes by employing the one-dimensional sodium dodecyl sulfate−polyacrylamide gel electrophoresis (1D SDS-PAGE) method, and 35 major protein bands obtained from three genotypes with contrasting seed protein profiles were further analyzed by mass spectrometry (MS). Twenty-five of these protein bands (MW range: ∼ 9−107 kDa) had significant ($p \leq 0.05$) matches to polypeptides in protein databases. MS analysis showed that most of the analyzed protein bands contained more than one protein type and, in total, over 100 proteins were identified. These included major seed storage proteins such as legumin, vicilin, and convicilin, as well as other protein classes like lipoxygenase, heat shock proteins, sucrose-binding proteins, albumin, and defensin. Furthermore, seed protein extracts were separated by size-exclusion high-performance liquid chromatography (SE-HPLC), and percentages of the major protein classes were determined. On average, legumin and vicilin/convicilin accounted for 50 and 27% of the total protein extract, respectively. However, the proportions of these proteins varied considerably among genotypes, with the ratio of legumin:vicilin/convicilin ranging from 1:1 to 1:3. In addition, there was a significant ($p < 0.01$) negative correlation between the contents of these major fractions ($r = -0.83$). This study significantly extends the number of identified *Vicia faba* seed proteins and reveals new qualitative and quantitative variation in seed protein composition, filling a significant gap in the literature. Moreover, the germplasm and screening methods presented here are expected to contribute in selecting varieties with improved protein content and quality.

KEYWORDS: *vicia faba, legumin, vicilin, protein quantification, SE-HPLC*

INTRODUCTION

Faba bean (*Vicia faba*, hereafter *Vf*) seeds contain about 29% protein,[1] and the crop is well adapted to various climates and is grown for both human and animal nutrition.[2,3] Given its high yield potential[4] and unparalleled nitrogen-fixation capacity,[5] *Vf* is among the few crops with great potential to meet the dietary needs of the growing human population while maintaining sustainability of agricultural production systems.[6] Much of the research on seed quality to date has focused on the reduction or removal of antinutrients, namely, vicine and convicine[7−9] and seed coat tannins,[10−12] with surprisingly little effort dedicated to improving the protein composition.

Utilization of plant proteins for human or animal nutrition is largely determined by the nutritional and functional properties of their constituent protein classes. It is estimated that *Vf* seed proteins contain ∼ 80% globulin which in turn is comprised of legumin and vicilin/convicilin, also referred to as 11S and 7S, respectively, based on their ultracentrifugation sedimentation coefficients,[13] respectively. Globulins belong to the cupin superfamily,[14] and the legumin and vicilin types have a high degree of structural homology.[15,16]

Legumin is a major *Vf* seed protein, representing about 50% of the storage proteins.[17,18] It is encoded by multiple genes belonging to type-A (Methionine-containing) and type-B (Methionine-lacking) subunits.[18] Only few genes encoding type-A (A1 and A2), type-B (LeB2, LeB4, LeB6, and LeB7), and one high-molecular mass legumin polypeptide (LeB3) have been described[18−20] in the literature. However, Tucci et al.[21] reported 29 biochemically distinct disulfide-linked $\alpha\beta$ legumin subunit pairs with molecular weights between 39−81 kDa, suggesting that the number of legumin-encoding genes could be much more than is currently known. Vicilin is also a heterogeneous protein in its native trimer form.[21] Regarding convicilin subunits, at least two structural genes have been described,[22] though the question of whether convicilin can be considered a vicilin subunit or a distinct class of globulin is yet to be resolved in *Vf*.

The relationship between subunit composition of major storage proteins and the overall seed protein quality has been studied in other legumes like soybean,[23] where molecular markers for specific legumin and vicilin-like subunit variants with desirable qualities have been developed.[24] In *Vf*, it is generally accepted that selection for a higher legumin:vicilin ratio could enhance its nutritional quality[1,25,26] since some major legumin subunits contain relatively higher proportions of sulfur-containing amino acids (S-AA) compared to vicilin. However, given the underlying genetic complexity of these broadly defined classes of storage proteins, concrete exploitation of genetic variation in seed protein composition

Received: May 8, 2020
Revised: July 14, 2020
Accepted: July 17, 2020
Published: July 17, 2020

Table 1. List of *Vf* Genotypes Used for Qualitative and Quantitative Analysis of Seed Protein Composition

genotype	original source	germplasm category	country[a]	mapping populations found[b]
LG Cartouche		cultivar	UK	
Lynx		cultivar	UK	
Vertigo		cultivar	UK	RSBP
Wizard		cultivar	UK	
Fanfare		cultivar		RSBP
Icarus	Icarus	inbred line from cultivar	Ecuador	7-way MAGIC; Icarus × Ascot
NV640	Maris Bead	inbred line from cultivar	UK	RSBP
NV643	Albus	inbred line from cultivar	Poland	Albus × BPL10; RSBP; 7-way MAGIC
NV672	Betty	inbred line from cultivar		RSBP
NV866	Disco/2	inbred line from cultivar	France	Hedin/2 × Disco/2; 4WP; RSBP
NV639−2	Hedin	inbred line from cultivar	Germany	RSBP
RV501	Robin Hood	inbred line from cultivar	UK	
RV502	Sutton	inbred line from cultivar	UK	
RV503	Casata Midwinter	inbred line from informal cultivar	UK	RSBP
RV504	Crimson Flowered-3	inbred line from heirloom cultivar	UK	RSBP
RV505	Diana	inbred line from cultivar	Canada	7-way MAGIC
RV506	Cuscan Super Yellow-1	partial Inbred line from landrace	Peru	RSBP
RV507	Iantos-3	partial inbred line from landrace	Peru	RSBP
RV508	Mustard Yellow	partial inbred line from landrace	Peru	RSBP
RV509	Sakha4	inbred line from cultivar	Egypt	RSBP
RV510	Nubaria3	inbred line from cultivar	Egypt	RSBP
RV511	Misr3	inbred line from cultivar	Egypt	RSBP
RV512	Giza716	inbred line from cultivar	Egypt	RSBP
NV735	Mélodie	inbred line from cultivar	France	Melodie×ILB938−2; RSBP
RV319−2		inbred line	UK	
NV153	ig12658	inbred line from Landrace	Ethiopia	
NV648−1	BPL10	inbred line	Unknown	Albus × BPL10; RSBP
NV734	ILB938−2	inbred line	Colombia	Melodie×ILB938−2; 4WP; 7-way MAGIC
NV657	INRA 29H	inbred line	France	RSBP
L170	ig132238	inbred line	China	4WP
NV651−3	BPL21	inbred line	Unknown	RSBP
NV658−2	CGN07715 cf-3	inbred line	Unknown	
L43	ig114476	inbred line	Bangladesh	4WP
RV322	HEL170	inbred line	China	RSBP
NV873−13	F5 from NV644xNV153	recombinant inbred line	Unknown	RSBP

[a]Country of release (for cultivars) or collection (landrace materials). [b]RSBP: Reading Spring Bean Population (currently under development); 7-way MAGIC: Multiparent advanced generation intercross (under development); 4WP: 4-Way cross population.[36]

for the development of cultivars with improved protein profiles would require identification of the genes encoding the major seed storage proteins, as well as understanding their synthesis, transport, and storage mechanisms. To date, studies have referred to just a few major protein subunits of legumin and vicilin[17,21,25−27] and although Liu et al.[28] identified several additional nonglobulin seed storage proteins from *Vf* by mass spectrometry, the identification of the full set of proteins that contribute to the nutritional and functional properties of the *Vf* seed is far from complete.

The one-dimensional sodium dodecyl sulfate−polyacrylamide gel electrophoresis (1D SDS-PAGE) method has been exploited in the qualitative and quantitative analysis of protein composition in various legume species.[24,29−31] However, the main problem inherent in this method is that depending on the particular electrophoresis conditions used, unrelated proteins of similar mobility can partially or completely overlap, which can lead to over- or underestimation of certain protein subunits. An alternative method of protein separation based on size-exclusion high-performance liquid chromatography (SE-HPLC) has been widely used in studying seed proteins of wheat, notably in determining the proportions of gliadin and glutenin fractions associated with certain quality attributes, including pasta-cooking and bread-making qualities.[32−35] The advantage of this method is that proteins can be quantified in their native condition and the sample analysis is amenable for automation.

In this study, our aim is to (1) assess the diversity in subunit composition of major *Vf* seed proteins in genetically diverse germplasm; (2) accurately identify the most abundant seed proteins; and (3) quantify the proportions of legumin and vicilin/convicilin proteins using a panel of diverse *Vf* genotypes.

MATERIALS AND METHODS

Reagents. Sodium phosphate, calcium chloride, trichloroacetic acid, dithiothreitol, iodoacetamide, triethylammonium bicarbonate, and a Bradford assay reagent were obtained from Sigma-Aldrich (UK). PageBlue, NuPAGE LDS sample buffer, and NuPAGE MES SDS running buffer were sourced from Thermo Fisher Scientific (UK). Acetonitrile, sulfuric acid, and HPLC grade water were obtained from Fisher Scientific (UK). A sequence grade porcine trypsin enzyme was obtained from Promega (UK).

Plant Materials. Thirty-five *Vf* genotypes, including inbred lines derived from breeding materials, landraces, and cultivars from

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

different locations around the world, were used in this study for 1D SDS-PAGE and SE-HPLC protein subunit profiling (Table 1). This genetically diverse population contained genotypes collected by the University of Reading (UK), the Agricultural Research Center (Egypt), Nordic Seeds (Denmark) and the University of Saskatchewan (Canada). The majority of these genotypes are parents of *Vf* mapping populations, which are already existing[36] or currently under development.

Total Protein Extraction. Five to 10 seeds per genotype were dried in an oven at 80 °C for 48 h and ground using a Laboratory Mill 3303 (Perten Instruments, Warrington, UK). The flour was then sieved through a 1 mm diameter sieve to obtain a homogenous sample. Total seed proteins were extracted according to the procedure reported by Mertens et al.[37] with some modifications. Briefly, we used 0.1 M phosphate buffer (pH 7.2) containing 5 g L^{-1} of potassium sulfate with a sample/buffer ratio of 1:10 (w/v). Samples were vortexed briefly and stirred for 30 min at 300 rpm followed by centrifugation at 20,000 × g for 30 min at room temperature. The supernatant was then transferred to a new tube and stored at −20 °C until further analysis. The protein concentration in protein extracts was measured using the Bradford method[38] with a SpectraMax i3x microplate reader (Molecular Devices, UK).

Protein Fractionation. The total seed protein extracts were fractionated by sequential extraction in aqueous and salt solutions to obtain fractions enriched for water-soluble and salt-soluble proteins (for details, see Figure S1). The globulin precipitation step was conducted according to the procedure reported by Krishnan et al.[39] A total of five protein fractions (hereafter F1−5) were obtained: water soluble (F1), globulin-depleted water soluble (F2), salt-soluble (F3), globulin-depleted salt-soluble (F4), and globulin-enriched fraction (F5). These fractions were then analyzed by SE-HPLC and SDS-PAGE.

Protein and Sulfur Content Analysis. Nitrogen and sulfur contents (%) were determined using an isotope ratio mass spectrometer (DELTA V IRMS, Thermo Fisher, UK). The analysis was carried out in duplicate using oven-dried flours of ∼ 1 mg. Nitrogen content data were then converted to protein content as: protein (%) = %N × 5.4.[40]

1D SDS-PAGE Analysis. One-dimensional SDS-PAGE analysis of the total protein extract (∼ 15 μg per well) was performed using NuPAGE 10% Bis−Tris precast gels. Before gel loading, the samples were mixed with the NuPAGE LDS sample buffer and a sample reducing agent following manufacturer's instructions. Gels were run in the NuPAGE MES SDS buffer in an XCell SureLock Mini-Cell at a constant current of 70 mA and a maximum voltage of 200 V for 1 h. Before staining, the gels were fixed with 12% trichloroacetic acid for 15 min and washed twice with 250 mL of deionized water for another 15 min on a rocker. Gels were then stained with 50 mL of a PageBlue protein-staining solution for 2 h followed by destaining overnight with deionized water.

Identification of Major Seed Protein Subunits. *In-Gel Protein Digestion.* Individual protein bands were carefully excised from gel lanes of the selected genotypes (LG Cartouche, NV657, and NV734) and were destained in 0.6 mL tubes with 400 μL of 50% acetonitrile (MeCN) and 50% 10 mM triethylammonium bicarbonate (TEAB) overnight. Gel pieces were then reduced with 10 mM dithiothreitol (DTT) in 10 mM TEAB for 30 min at 50 °C followed by alkylation with 50 mM iodoacetamide in 10 mM TEAB for 30 min in the dark. After washing three times with 400 μL of 10 mM TEAB and once with MeCN, the dehydrated gel samples were resuspended in 10 μL of 10 mM TEAB containing 200 ng of porcine trypsin and incubated at 25 °C overnight. The gel digests were placed on dry ice for 5 min, then allowed to thaw, and 30 μL of 10% MeCN/5% formic acid was added. After 15 min of sonication, peptide extracts were transferred to 250 μL PCR tubes. This step was repeated twice, and the resultant extract was pooled and dried in a centrifugal vacuum concentrator.

Mass Spectrometry Analysis. The dried peptides were resuspended in 20 μL of LC−MS buffer A (0.1% formic acid in water), and 10 μL of the sample was injected into an Ace C18 column (150 × 2.1 mm, 5 μM particle size with 300 Å pore size) and analyzed by LC−MS using a Thermo Scientific LTQ-Orbitrap XL interfaced with an Accela HPLC instrument. Buffer B was 0.1% formic acid in MeCN. The gradient was as follows: 0−2 min; 5% B, 20 min; 60% B, 20.1−23 min; 80% B, 23.1−30 min; 5% B. The column oven was maintained at 30 °C, and at 15 °C for the autosampler. The first 2 min and the last 6 min of each run were excluded from the analysis. A data-dependent acquisition (DDA) strategy was employed. In brief, ions were measured using the Orbitrap at 30,000 resolution, scanning from 400−2000 m/z. Three ions from each MS1 scan that were most abundant and multiply charged were chosen for MS2. MS2 was performed using collision-induced disassociation (CID) in the ion trap and scanned out at a unit resolution. The acquired data were analyzed using an in-house version of MASCOT search engine (Matrix Science, UK) via Mascot Daemon with file conversion performed using ProteoWizard. The acquired MS spectra were searched against the NCBI nonredundant protein (https://www.ncbi.nlm.nih.gov), common Repository of Adventitious Proteins (ftp://ftp.thegpm.org/fasta/cRAP), and other contaminant databases. The search parameters were set as follows: type of search = MS/MS ion search, enzyme = trypsin, variable modifications = acetyl (protein N-term), carbamidomethyl (C), Gln- > pyro-Glu (N-term Q), oxidation (M), mass values = monoisotopic, protein mass = unrestricted, peptide mass tolerance = ±10 ppm, fragment mass tolerance: ±1 Da, max missed cleavages = 2, and instrument type: ESI-TRAP.

Protein Composition Analysis by SE-HPLC. Size-exclusion HPLC analysis was conducted with the Waters Alliance 2695 Separations Module using a Phenomenex BioSep-SEC-S2000 column with silica resin (300 × 7.8 mm, 5 μm particle size and 145 Å pore size). The same extraction buffer (0.1 M phosphate buffer containing 5 g L^{-1} of potassium sulfate, pH = 7.2) was used as a mobile phase with a flow rate of 0.5 mL min^{-1}. The injection volume of protein sample was 20 μL, and detection was at 210 nm using the Waters 2996 photodiode array (PDA) detector. Two biological replicates were analyzed for each genotype, and the raw chromatogram data were exported for peak integration in Origin software (OriginLab Corporation, Northampton, MA, USA).

RESULTS AND DISCUSSION

A Comprehensive Survey of *Vf* Seed Proteins. In order to capture the most common seed protein variants, we first conducted a preliminary 1D SDS-PAGE screening of 35 diverse genotypes for their seed protein profiles (Figure S2). From this analysis, we identified three genotypes—LG Cartouche, NV657, and NV734—with distinct protein profiles (Figure 1) and used them for protein band identification. Forty-six bands, with apparent molecular weights (MW) ranging from less than 10 to ∼ 145 kDa on reducing 1D SDS-PAGE gels, were detected collectively from these three genotypes. Thirty-five of these bands were excised from the gel and subjected to mass spectrometry analysis, with 25 of them reporting significant ($p \leq 0.05$) matches with proteins in the database, mainly from *Vf* and related legumes (Table 2). Failure to identify the remaining 10 bands can be explained in term of their relatively lower abundance, which made it technically challenging to elute enough protein.

Nearly all analyzed bands contained more than one type of protein, and a total of 106 proteins were identified (for detailed list, see Table S1). As expected, the most abundant proteins were globulins, with polypeptides belonging to legumin, vicilin, and convicilin identified in 13, 8, and 4 of the 25 bands, respectively (Table 2). This wide molecular mass distribution of legumin and vicilin subunits was previously reported[21] using antibodies specific to these proteins. However, in the case of convicilin, for which a single discrete band near 68 kDa has been so far reported in the literature,[1,17,21,28] we have identified multiple bands, including a major band at ∼ 54 kDa (Figure 1,

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

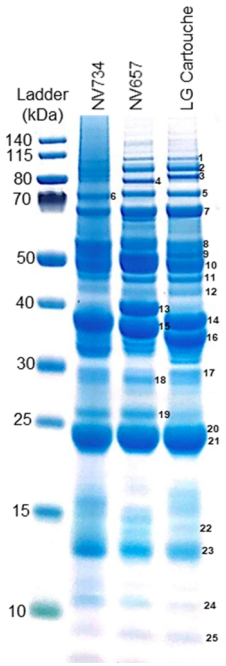

Figure 1. SDS-PAGE profile of three *Vf* genotypes with distinct seed protein profiles which were used for seed protein identification by mass spectrometry analysis.

Table 2). Although this is a new observation for *Vf*, it is not surprising considering that multiple convicilin subunits with MW ranging from 52−99 kDa have been reported in the related species *Medicago truncatula*.[31,41] To further investigate whether the two major convicilin bands identified (7, 8 in Figure 1) represent the convicilin A and B genes reported in the past,[22] we compared the protein sequences derived from these convicilin genes with the MS peptide sequences from band 7 and 8 of LG Cartouche and ILB 938−2. While convicilin B-specific peptides were found in both bands 7 and 8 in both genotypes, three peptides unique to convicilin A were found only in band 8 of LG Cartouche (Table S2). Furthermore, peptide sequences from band 7 of LG Cartouche contained a 37 AA long peptide which aligned to a region with significant polymorphism between convicilin A and B. Interestingly, this unique peptide had nine and five mismatches with A and B genes, respectively, but had 100% similarity with a convicilin accession (CAP06324.1) from *Lathyrus ochrus*. Taken together, these results indicate that convicilin structural diversity in *Vf* is greater than previously thought, comprising of at least two B-type isoforms, as well as A and other unnamed convicilin polypeptides which appear to be expressed in a genotype-dependent manner.

Mass spectrometry analysis also identified several less abundant but nonetheless distinct and well-conserved protein bands. These include two distinct lipoxygenase bands (~ 96 and 106 kDa), a heat shock protein (~ 73−75 kDa, depending on the genotype), a sucrose-binding protein (~ 45 kDa), albumins (~ 10.2, 12.4 and 13.7 kDa), and defensins (Table 2, Figure 1). From a nutritional quality point of view, lipoxygenase is considered antinutritional due to its role in lipid oxidation, which also leads to undesirable flavors during food processing.[42] The studied genotypes show noticeable variation in the intensity of lipoxygenase bands (Figure S2), but establishing the significance of this variation requires further scrutiny. In other legumes such as soybean[43] and pea,[44] efforts to develop genotypes lacking the major seed protein lipoxygenase have been successful.

Protein Subunit Diversity Among *Vf* Genotypes. In total, we identified 15 protein bands polymorphic among the *Vf* genotypes, with variation being concentrated in less abundant proteins with MW of more than 70 kDa or less than 20 kDa (Figure S2). The most interesting protein variants were found in the α subunits of legumin, represented in the majority of *Vf* genotypes by a single legumin band of about 38 kDa and by rare legumin α subunits of about 36 and 40 kDa in LG Cartouche and NV657, respectively (Figure 1). MS analysis showed that the higher MW legumin α subunit in NV657 is an A-type legumin while that of a lower mass in LG Cartouche is a B-type legumin α subunit (Table 2). Further evidence that these genotypes contain novel legumin subunits comes from the observation that unreduced proteins of these genotypes have two distinct major bands of $\alpha\beta$ polypeptides (data not shown). These natural variants in subunit composition can be exploited to address questions on the genetic architecture of seed protein composition and the impact of discrete protein subunit variants on the nutritional and processing quality of the overall seed protein.

Although the majority of the analyzed bands contained one predominant protein type, the existence of some bands where there is an overlap between major bands of different protein classes underpins the need for an alternative method to the conventional SDS-PAGE-based densitometric approach for quantifying protein composition. Nonetheless, this expanded and refined list of identified seed proteins can be utilized as a reference for qualitative SDS-PAGE-based screening for protein subunit variants of interest in breeding and research materials like mutant or mapping populations.

SE-HPLC Analysis of Seed Proteins. *Total Seed Protein Extract.* The total seed protein extract from the NV639−2 inbred line was separated using a Phenomenex BioSep-SEC-S2000 column, producing chromatographic peaks between 10 and 28 min of the analysis time (Figure 2), and four major peaks (1, 2, 4, and 18) accounted for more than 70% of the total chromatogram peak area. To confirm the identity of proteins associated with these peaks, SE-HPLC peak fractions were collected at 1 min intervals and separated by 1D SDS-PAGE. By comparing these gels with the annotated SDS-PAGE (on the basis of MS analysis), it was determined that peaks 2 and 4 were legumin and vicilin/convicilin aggregates with retention times of 12.4 and 14.0 min, respectively (Figure 2). Proteins with smaller molecular weights were eluted in the expected order, suggesting that the selected column was suitable for the separation of *Vf* proteins. However, despite having strong signals at 214, 254, and 280 nm, no detectable proteins were found in peak 1 and all other peaks eluted after ~ 21 min (Figure 2). We therefore hypothesized that peak 1 corresponds to protein−phenol complexes that could not be detected by SDS-PAGE. Sęczyk et al.[45] found that some phenolic compounds preferentially interact with globulins, leading to changes in their SE-HPLC and SDS-PAGE profiles. Regarding peak 18, Defaix et al.,[46] who used the same type of column used in our study, suggested that the strong signal near the end of the analysis was due to phenolic compounds. To further investigate this hypothesis, SE-HPLC profiles of protein extracts of dehulled and whole seeds were compared; since *Vf* seed coats contain higher phenols, we would expect the proportion of peak 1 and 18 to be substantially reduced in the dehulled sample. Indeed, dehulled protein samples showed

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

Table 2. Major Proteins Identified by Mass Spectrometry Analysis of Protein Bands Excised from Reducing SDS-PAGE Gels of *Vf* Seed Proteins and their Significant ($p \leq 0.05$) Matches (from *Vf* and Other Legume Species) in the Database

SDS-PAGE band[a]	band apparent MW (kDa)	accession	score	num. of significant sequences	emPAI	description	species
1	106.9	gil126405	565	15	1.08	seed linoleate 9S-lipoxygenase-3	*Pisum sativum*
		gil 164512572	128	2	0.18	convicilin	*Vf*
2	96.3	gil126405	508	15	1.18	seed linoleate 9S-lipoxygenase-3	*Pisum sativum*
		gil 164512572	178	4	0.39	convicilin	*Vf*
3	88.8	gil 164512572	120	2	0.18	convicilin	*Vf*
		gil 187766747	99	1	0.26	Gly m Bd 28 K allergen	*Glycine max*
4	83	gil 164512572	165	6	0.68	convicilin	*Vf*
		gil22053	154	9	1.34	vicilin: precursor	*Vf*
5	75.2	gil 357480003	391	8	0.81	heat shock 70 kDa protein	*Medicago truncatula*
		gil126162	94	4	0.74	legumin type B	*Vf*
6	73.1	gil562006	364	12	1.26	PsHSP71.2	*Pisum sativum*
		gil 164512572	123	4	0.40	convicilin	*Vf*
7	64.7	gil 164512572	1145	25	6.89	convicilin	*Vf*
		gil126164	101	3	0.30	legumin type B: precursor	*Vf*
8	54.1	gil 164512572	1074	21	4.67	convicilin	*Vf*
		gil403336	312	7	0.68	legumin-related high-molecular weight polypeptide	*Vf*
		gil3122060	123	6	0.78	elongation factor 1-alpha	*Vf*
9	50.0	gil137584	1344	22	6.28	vicilin: precursor	*Vf*
		gil403336	589	11	1.25	legumin-related high-molecular weight polypeptide	*Vf*
10	48.2	gil137584	1374	22	6.28	vicilin: precursor	*Vf*
		gil403336	342	7	0.68	legumin-related high-molecular weight polypeptide	*Vf*
		gil12580894	176	6	0.69	putative sucrose-binding protein	*Vf*
11	45.4	gil12580894	1018	18	4.40	putative sucrose-binding protein	*Vf*
		gil22008	226	9	1.16	legumin A2 primary translation product	*Vf*
		gil126166	178	8	1.84	legumin type B	*Vf*
12	43.5	gil2578438	98	3	0.26	legumin (minor small)	*Pisum sativum*
		gil403336	90	3	0.26	legumin-related high-molecular weight polypeptide	*Vf*
13	40.2	gil22008	662	14	2.51	legumin A2 primary translation product	*Vf*
		gil 164512572	208	7	0.78	convicilin	*Vf*
		gil259474	312	6	1.42	legumin propolypeptide alpha chain	*Vf*
14	38.4	gil22008	875	14	2.61	legumin A2 primary translation product	*Vf*
		gil126166	628	12	3.78	legumin type B	*Vf*
		gil22053	392	11	1.75	vicilin: Precursor	*Vf*
15	37.6	gil542002	823	9	2.67	legumin type B alpha chain; precursor	*Vf*
		gil137584	506	16	3.24	vicilin: precursor	*Vf*
		gil22008	312	10	1.31	legumin A2 primary translation product	*Vf*
16	36.2	gil542002	926	8	2.28	legumin type B alpha chain: precursor	*Vf*
		gil137584	747	19	4.83	vicilin: precursor	*Vf*
		gil22008	253	7	0.83	legumin A2 primary translation product	*Vf*
17	31.5	gil137584	277	11	1.71	vicilin: precursor	*Vf*
		gil137582	203	4	0.44	vicilin: precursor	*Vf*
18	30.4	gil137584	300	11	1.73	vicilin: precursor	*Vf*
		gil137582	157	4	0.45	vicilin: precursor	*Vf*
19	26.0	gil22008	76	2	0.18	legumin A2 primary translation product	*Vf*
		gil29539109	54	3	0.35	allergen len c	*Lens culinaris*
20	24.0	gil12580894	53	1	0.09	putative sucrose-binding protein	*Vf*
21	22.3	gil259475	399	5		legumin propolypeptide beta chain	*Vf*
		gil403336	369	5		legumin-related high-molecular weight polypeptide	*Vf*

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

Table 2. continued

SDS-PAGE band[a]	band apparent MW (kDa)	accession	score	num. of significant sequences	emPAI	description	species
22	13.7	gil51704211	97	2	0.98	albumin-1 E	*Pisum sativum*
23	12.4	gil51704211	72	1	0.27	albumin-1 E	*Pisum sativum*
		gil27466894	70	2	0.68	thioredoxin h	*Pisum sativum*
		gil 763805274	50	1	0.25	hypothetical protein	*Gossypium raimondii*
24	10.2	gil51704209	60	1	0.29	albumin-1 C	*Pisum sativum*
25	9.5	gil 205277584	56	2	1.15	defensin-like protein	*Vf*
		gil 205277582	55	2	1.19	defensin-like protein	*Vf*

[a]Band numbers in the first column refer to the band numbers shown in Figure 2.

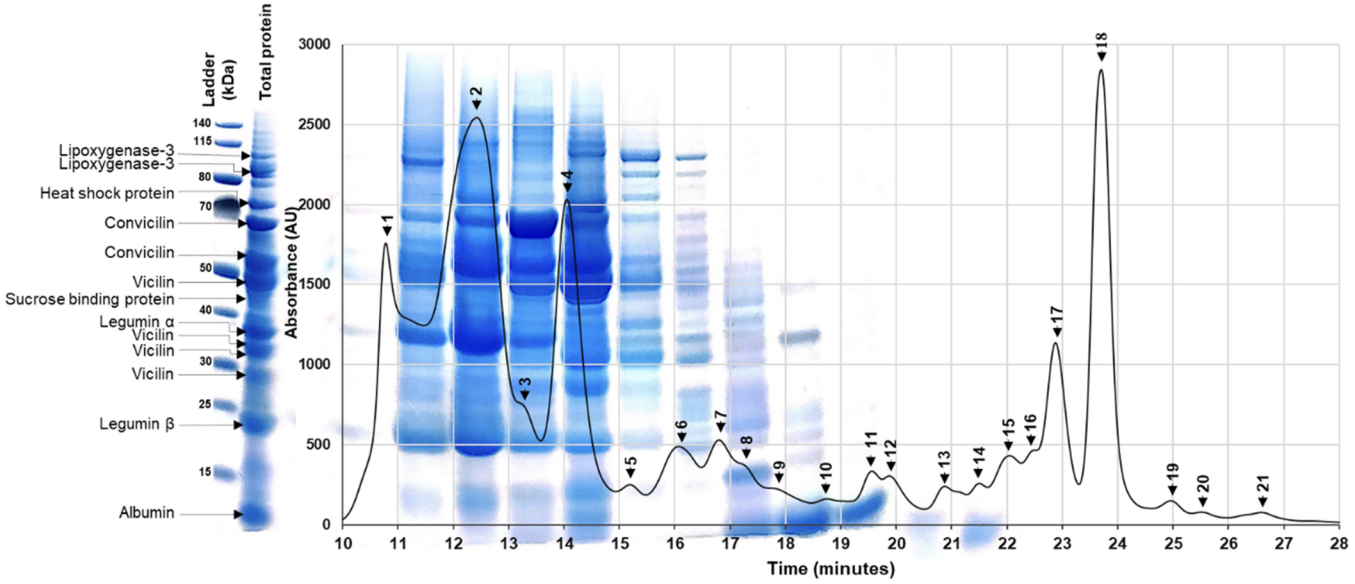

Figure 2. SE-HPLC chromatogram of *Vf* seed protein extract from NV639−2 which is overlaid with the SDS-PAGE profile of protein fractions collected at a 1 min interval across the analysis time. Observable peaks are numbered from 1−21 and labels on the left refer to some of the major protein subunits identified in this study.

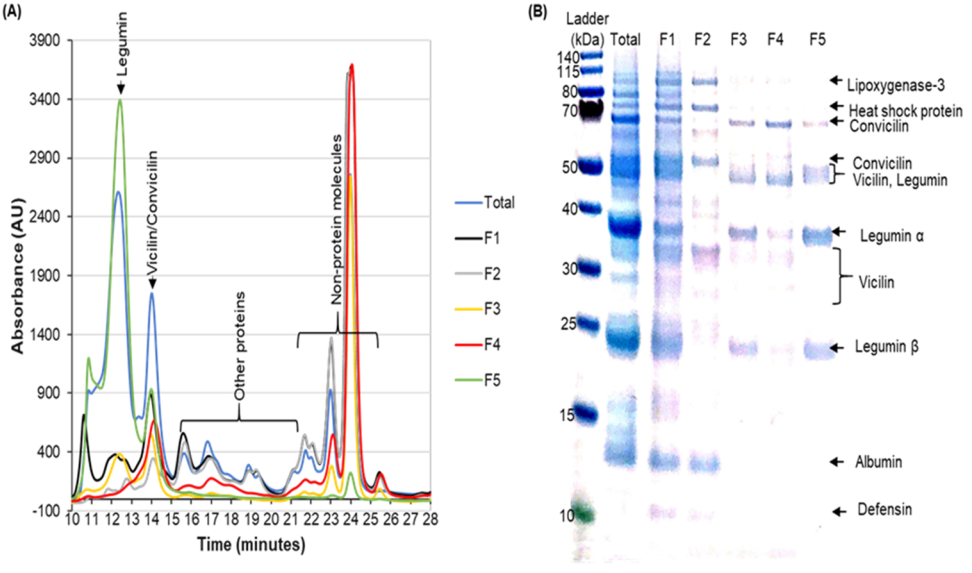

Figure 3. (A) SE-HPLC chromatogram and (B) SDS-PAGE profile of fractionated proteins of NV639−2 line. Fractions (F1−5) are water-extractable proteins (F1), globulin-removed water-soluble fraction by addition of 10 mM $CaCl_2$ (F2), pellet from F1 extracted with 0.1 mM phosphate buffer at pH = 7.2 (F3), globulin-depleted salt-soluble proteins by addition of 10 mM $CaCl_2$ (F4), and pellet from F4 suspended in 0.1 mM phosphate buffer at pH = 7.2 (F5).

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

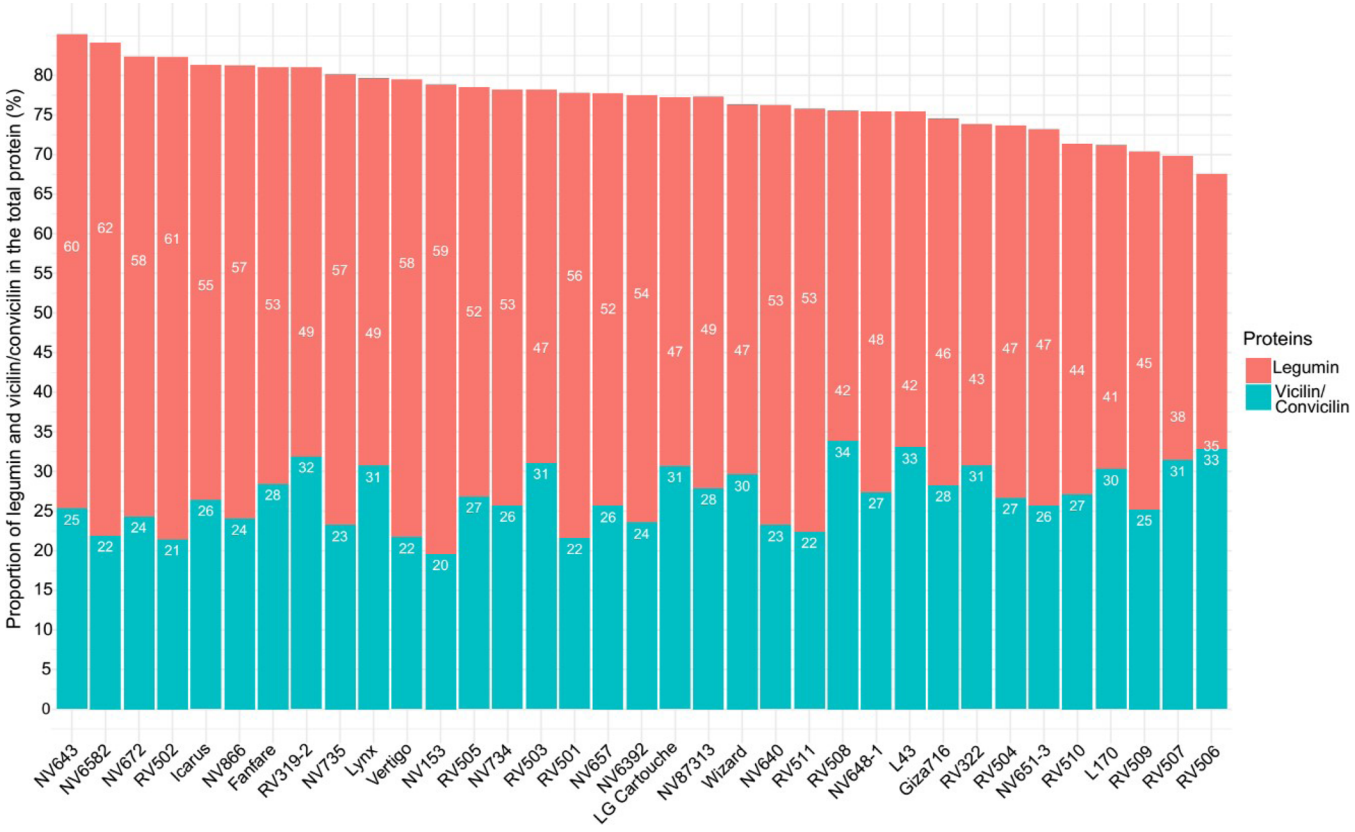

Figure 4. Bar graph showing the proportions of legumin and vicilin/convicilin in the total seed protein extracts of 35 *Vf* genotypes. Protein percentages are determined from the relative area of SE-HPLC peaks belonging to each protein class in two biological replicates.

a nearly 50% decrease in both peaks (Figure S3). On this basis, only the peak area between 11.5 and 21.5 min was considered for SE-HPLC protein composition analysis.

Fractionated Seed Proteins. To further confirm that the peaks resolved by SE-HPLC belong to the major seed proteins, we separated the protein fractions prepared by sequential extraction (denoted as F1−F5 in Figure 3) both by SE-HPLC and 1D SDS-PAGE. Comparison of the separation profiles obtained for F1−F5 protein fractions in the two systems (Figure 3A,B) showed that the functional proteins like lipoxygenase, heat shock protein, and albumin have relatively higher solubility in water and they were enriched in F1 and F2, with an elution time between 15 and 20 min under the SE-HPLC conditions used in this work. However, since these peaks, unlike globulin peaks, were poorly resolved by SE-HPLC, they are referred collectively as "other proteins" as shown in Figure 3A. On the other hand, different globulin subclasses were enriched in specific extraction buffers, reflecting their different physicochemical properties. For instance, legumin is soluble in water and to a higher degree in a phosphate buffer (F3), which could be further enriched by the addition of Ca^{+2} (F5), producing a major HPLC peak with a retention time of 12.4 min. However, the vicilin/convicilin subclass of globulin appears to contain a mixture of subunits with varying properties. As shown in Figure 3B, protein bands corresponding to subunits of convicilin (~ 54 kDa) and vicilin (~ 37 kDa) were extractable in water (F1) and did not precipitate in the presence of Ca^{+2} (F2). Conversely, other subunits of convicilin (~ 65 kDa) and vicilin (~ 50 kDa) were soluble in a phosphate buffer (F3) and precipitated, to a certain degree, with the addition of Ca^{+2} (F4 and F5). Even though the fractions F1, 4, and 5 have a vicilin/convicilin peak of nearly a similar magnitude, the SDS-PAGE profile of these different fractions showed distinct subsets of vicilin/convicilin (Figure 3B). It was therefore concluded that convicilin and vicilin polypeptides form heterogeneous subclasses of the globulin type protein with distinct physicochemical properties but eluted as a single peak with a retention time of 14 min under the SE-HPLC conditions used in our study. This observation would explain why O'Kane et al.,[47] who conducted various fractionation and physicochemical characterization of vicilin and convicilin proteins in pea, concluded that convicilin is a α subunit of vicilin.

Quantification of Legumin and Vicilin/Convicilin Contents by SE-HPLC. Since one of the major indicators of protein quality is the content of S-AA which in turn is determined by the relative proportions of the major protein classes, the SE-HPLC method was used to quantify legumin and vicilin/convicilin contents in a panel of 35 genetically diverse *Vf* genotypes. Overall, legumin and vicilin/convicilin accounted for 50 and 27% of the protein extract, respectively. Among the genotypes, legumin accounted for 35 to 62% of the quantified peak area while vicilin/convicilin for 20−34% (Figure 4). These results are comparable with the findings of Utsumi, Yokoyama, and Mori[27] who reported ranges of 42 to 47% and 28 to 31% for 11 and 7S globulins in crude protein extracts of six *Vf* cultivars analyzed by the sucrose density gradient fractionation technique. In another study, the *Vf* legumin and vicilin content reportedly varied between 40 to 45% and 20 to 25%, respectively.[3] Moreover, according to our study, globulin peaks represent 77% of the total protein peak area, which is very close to the estimated 70 to 80% globulin content in *Vf*

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

seed proteins reported by other studies.[3,17,48] The present study appears to capture a wider variation in *Vf* protein composition than previously reported, likely reflecting the fact that the plant materials we used spanned a deliberately broad genetic base.

The legumin to vicilin/convicilin ratio (L/V) varied from 1 to 3, which is comparable to the 2.1 to 3.6 range reported previously.[26] Among the genotypes of special interest for their high L/V ratio are two inbred lines, NV153 and NV658−2, which have been previously used as parents in mapping populations (Table 2). However, it is important to mention that this ratio is highly sensitive and can be affected by many factors, including genotype, environment, protein extraction method, and quantification techniques. In this work, the reproducibility of the results was measured by comparing five replicates of a single genotype (NV639−2) that were independently extracted and analyzed in different batches. The coefficient of variation between the five replicates was higher in the legumin fraction (6%) compared to vicilin/convicilin (3%). However, biological replicates of each genotype analyzed in the same run were highly correlated ($r^2 > 0.98$) (Figure S4). This indicates the importance of including the batch as a cofactor for statistical analysis.

Finally, we exploited this quantitative data from a wide spectrum of germplasm to examine the possible limits and trade-offs between the two main classes of storage proteins and overall sulfur and protein contents. The Pearson's correlation analysis showed that the legumin content significantly but negatively correlated with vicilin/convicilin ($r = -0.83$, $p < 0.001$) and with "other proteins" ($r = -0.87$, $p < 0.001$) (Figure 5). On the other hand, the seed sulfur content correlated negatively with the content of vicilin/convicilin, the S-AA-deficient fraction[1] but correlated positively with the total protein content. Interestingly, the total protein content was independent of the proportion of major protein fractions, suggesting that protein composition can be improved without compromising the protein content. Similar independence of the total protein and globulin fractions has been observed in pea,[30] while, in contrast, a highly significant negative correlation between the content of certain 7S fractions and the total seed protein content has been reported in soybeans.[35]

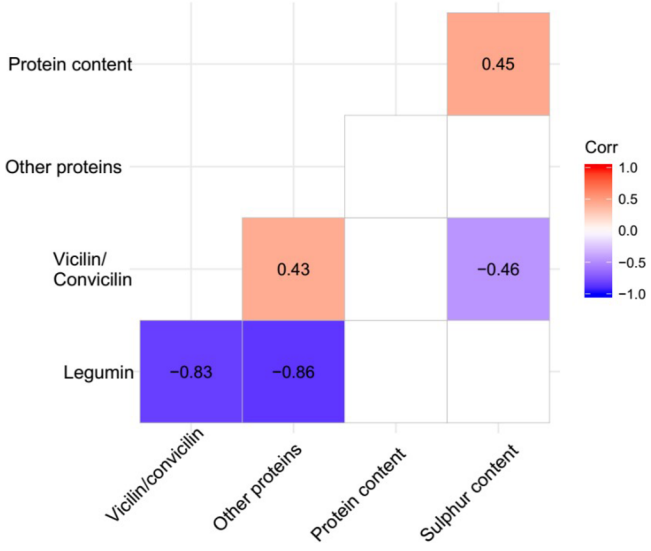

Figure 5. Correlation matrix between proportion of globulin fractions and other seed composition parameters at a significance level of $p \leq 0.05$.

In conclusion, this study provides a contemporary survey of the major seed proteins and their subunit composition among genetically diverse *Vf* germplasm. The MS-based identification of many major protein bands is a timely update linking a greater diversity of seed storage protein sequences to specific protein subunits which can be readily resolved with 1D SDS-PAGE gels; this new information can be used in the screening of germplasm with unique protein profiles, such as naturally occurring or induced mutations related to the reduced content of undesirable or increased content of desirable proteins. Also, we have demonstrated the potential of SE-HPLC as a method to efficiently determine the contents of legumin and vicilin/convicilin in *Vf*. This work paves the way for further understanding the *Vf* seed protein composition and the development of cultivars with desired protein quality.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jafc.0c02927.

Summary of the procedure used to fractionate *Vf* seed proteins based on their solubility in aqueous and salt solutions (Figure S1); protein subunit profiles of the 35 *Vf* genotypes separated by one-dimensional SDS-PAGE using 10% polyacrylamide gels (Figure S2); comparison between whole and dehulled *Vf* seeds for the proportion of two SE-HPLC peaks (1 & 18) (Figure S3); and correlation between the values of two biological replicates in the quantification of legumin and vicilin/convicilin by SE-HPLC (Figure S4) (PDF)

Detailed list of proteins identified by mass spectrometry analysis of protein bands from the seeds of three *Vf* genotypes (Table S1) and unique peptides belonging to convicilin genes (A & B) identified in major convicilin bands 7 and 8 shown in Figure 1 (Table S2) (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Paola Tosi − *School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AH, United Kingdom*; Phone: +44 (0) 118 378 8119; Email: p.tosi@reading.ac.uk

Authors

Ahmed O. Warsame − *School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AH, United Kingdom*; orcid.org/0000-0002-9281-0443

Nicholas Michael − *School of Chemistry, Food and Pharmacy, University of Reading, Reading RG6 6UR, United Kingdom*

Donal M. O'Sullivan − *School of Agriculture, Policy and Development, University of Reading, Reading RG6 6AH, United Kingdom*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jafc.0c02927

Funding

The authors thank the Islamic Development Bank (IsDB) for financial support to AOW during the preparation of this manuscript.

Notes

The authors declare no competing financial interest.

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

■ ACKNOWLEDGMENTS

The authors thank Chris Humphrey for helping with SE-HPLC analysis and Dr. Walid El-Rodeny (Agricultural Research Center, Egypt), Hamid Khazaei (University of Saskatchewan), and Jens Knudsen (Nordic Seeds) for providing seeds.

■ ABBREVIATIONS USED

S-AA, sulfur-containing amino acid; 1D SD-PAGE, one-dimensional sodium dodecyl sulfate−polyacrylamide gel electrophoresis; SE-HPLC, size-exclusion high-performance liquid chromatography; kDa, kilo dalton; emPAI, exponentially modified protein abundance index; MW, molecular weight

■ REFERENCES

(1) Warsame, A. O.; O'Sullivan, D. M.; Tosi, P. Seed Storage Proteins of Faba Bean (Vicia faba L): Current Status and Prospects for Genetic Improvement. *J. Agric. Food Chem.* **2018**, *66*, 12617−12626.

(2) Duc, G.; Aleksić, J. M.; Marget, P.; Mikic, A.; Paull, J.; Redden, R. J.; Sass, O.; Stoddard, F. L.; Vandenberg, A.; Vishnyakova, M.; Torres, A. M., Faba Bean. In *Grain Legumes*, Ron, A. M. D., Ed. Springer Science+Business Media: New York, 2015; Vol. *10*, pp 141−178.

(3) Multari, S.; Stewart, D.; Russell, W. R. Potential of Fava Bean as Future Protein Supply to Partially Replace Meat Intake in the Human Diet. *Compr. Rev. Food Sci. Food Saf.* **2015**, *14*, 511−522.

(4) Cernay, C.; Pelzer, E.; Makowski, D. A global experimental dataset for assessing grain legume production. *Sci. Data* **2016**, *3*, 160084.

(5) Baddeley, J. A.; Jones, S.; Topp, C. F. E.; Watson, C. A.; Helming, J.; Stoddard, F. L. *Biological nitrogen fixation (BNF) by legume crops in Europe*; 2013.

(6) Foyer, C. H.; Lam, H. M.; Nguyen, H. T.; Siddique, K. H.; Varshney, R. K.; Colmer, T. D.; Cowling, W.; Bramley, H.; Mori, T. A.; Hodgson, J. M.; Cooper, J. W.; Miller, A. J.; Kunert, K.; Vorster, J.; Cullis, C.; Ozga, J. A.; Wahlqvist, M. L.; Liang, Y.; Shou, H.; Shi, K.; Yu, J.; Fodor, N.; Kaiser, B. N.; Wong, F. L.; Valliyodan, B.; Considine, M. J. Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants* **2016**, *2*, 16112.

(7) Khazaei, H.; Purves, R. W.; Hughes, J.; Link, W.; O'Sullivan, D. M.; Schulman, A. H.; Björnsdotter, E.; Geu-Flores, F.; Nadzieja, M.; Andersen, S. U.; Stougaard, J.; Vandenberg, A.; Stoddard, F. L. Eliminating vicine and convicine, the main anti-nutritional factors restricting faba bean usage. *Trends Food Sci. Technol.* **2019**, *91*, 549−556.

(8) Khazaei, H.; Purves, R. W.; Song, M.; Stonehouse, R.; Bett, K. E.; Stoddard, F. L.; Vandenberg, A. Development and validation of a robust, breeder-friendly molecular marker for the vc-locus in faba bean. *Mol. Breed.* **2017**, *37*, 140.

(9) Khazaei, H.; O'Sullivan, D. M.; Jones, H.; Pitts, N.; Sillanpää, M.; Pärssinen, P.; Manninen, O.; Stoddard, F. L. Flanking SNP markers for vicine−convicine concentration in faba bean (Vicia faba L.). *Mol. Breed.* **2015**, *35*, 38.

(10) Gutierrez, N.; Avila, C. M.; Moreno, M. T.; Torres, A. M. Development of SCAR markers linked to zt-2, one of the genes controlling absence of tannins in faba bean. *Aust. J. Agric. Res.* **2008**, *59*, 62−68.

(11) Webb, A.; Cottage, A.; Wood, T.; Khamassi, K.; Hobbs, D.; Gostkiewicz, K.; White, M.; Khazaei, H.; Ali, M.; Street, D.; Duc, G.; Stoddard, F. L.; Maalouf, F.; Ogbonnaya, F. C.; Link, W.; Thomas, J.; O'Sullivan, D. M. A SNP-based consensus genetic map for synteny-based trait targeting in faba bean (Vicia faba L.). *Plant Biotechnol. J.* **2016**, *14*, 177−185.

(12) Zanotto, S.; Vandenberg, A.; Khazaei, H. Development and validation of a robust KASP marker for zt2 locus in faba bean (Vicia faba). *Plant Breed.* **2019**, 1−380.

(13) Chakraborty, P.; Sosulski, F.; Bose, A. Ultracentrifugation of salt-soluble proteins in ten legume species. *J. Sci. Food Agric.* **1979**, *30*, 766−771.

(14) Dunwell, J. M.; Purvis, A.; Khuri, S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **2004**, *65*, 7−17.

(15) Kesari, P.; Sharma, A.; Katiki, M.; Kumar, P.; Gurjar, B. R.; Tomar, S.; Sharma, A. K.; Kumar, P. Structural, Functional and Evolutionary Aspects of Seed Globulins. *Protein Pept. Lett.* **2017**, *24*, 267−277.

(16) Fukushima, D. Structures of plant storage proteins and their functions. *Food Rev. Int.* **1991**, *7*, 353−381.

(17) Müntz, K.; Horstmann, C.; Schlesier, B., Vicia globulins. In *Seed Proteins*, Shewry, P. R.; Casey, R., Eds. Springer Netherlands: Dordrecht, 1999; pp 259−284, DOI: 10.1007/978-94-011-4431-5_12.

(18) Horstmann, C.; Schlesier, B.; Otto, A.; Kostka, S.; Müntz, K. Polymorphism of legumin subunits from field bean (Vicia faba L. var. minor) and its relation to the corresponding multigene family. *Theor. Appl. Genet.* **1993**, *86*, 867−874.

(19) Fuchs, J.; Schubert, I. Localization of seed protein genes on metaphase chromosomes of Vicia faba via fluorescence in situ hybridization. *Chromosome Res.* **1995**, *3*, 94−100.

(20) Baumlein, H.; Wobus, U.; Pustell, J.; Kafatos, F. C. The legumin gene family: structure of a B type gene of Vicia faba and a possible legumin gene specific regulatory element. *Nucleic Acids Res.* **1986**, *14*, 2707−2720.

(21) Tucci, M.; Capparelli, R.; Costa, A.; Rao, R. Molecular heterogeneity and genetics of Vicia faba seed storage proteins. *Theor. Appl. Genet.* **1991**, *81*, 50−58.

(22) Sáenz de Miera, L. E.; Ramos, J.; Pérez de la Vega, M. A comparative study of convicilin storage protein gene sequences in species of the tribe Vicieae. *Genome* **2008**, *51*, 511−523.

(23) Poysa, V.; Woodrow, L.; Yu, K. Effect of soy protein subunit composition on tofu quality. *Food Res. Int.* **2006**, *39*, 309−317.

(24) Boehm, J. D.; Nguyen, V.; Tashiro, R. M.; Anderson, D.; Shi, C.; Wu, X.; Woodrow, L.; Yu, K.; Cui, Y.; Li, Z. Genetic mapping and validation of the loci controlling 7S α' and 11S A-type storage protein subunits in soybean [Glycine max (L.) Merr.]. *Theor. Appl. Genet.* **2018**, *131*, 659−671.

(25) Martensson, P., Variation in legumin: vicilin ratio between and within cultivars of Vicia faba L. var. minor. In Martinus Nijhoff. *World crops: production, utilization and description*, *3*.: The Hague, 1980; pp 159−172.

(26) Gatehouse, J.; Croy, R.; McIntosh, R.; Paul, C.; Boulter, D., Quantitative and qualitative variation in the storage proteins of material from the EEC joint field bean test. *Quantitative and qualitative variation in the storage proteins of material from the EEC joint field bean test*. 1980, 173−188.

(27) Utsumi, S.; Yokoyama, Z.-I.; Mori, T. Comparative Studies of Subunit Compositions of Legumins from Various Cultivars of Vicia faba L Seed. *Agric. Biol. Chem.* **1980**, *44*, 595−601.

(28) Liu, Y.; Wu, X.; Hou, W.; Li, P.; Sha, W.; Tian, Y. Structure and function of seed storage proteins in faba bean (Vicia faba L.). *3 Biotech.* **2017**, *7*, 74.

(29) Panthee, D. R.; Kwanyuen, P.; Sams, C. E.; West, D. R.; Saxton, A. M.; Pantalone, V. R. Quantitative trait loci for β-conglycinin (7S) and glycinin (11S) fractions of soybean storage protein. *J. Am. Oil Chem. Soc.* **2004**, *81*, 1005−1012.

(30) Tzitzikas, E. N.; Vincken, J.-P.; de Groot, J.; Gruppen, H.; Visser, R. G. F. Genetic Variation in Pea Seed Globulin Composition. *J. Agric. Food Chem.* **2006**, *54*, 425−433.

(31) Le Signor, C.; Aimé, D.; Bordat, A.; Belghazi, M.; Labas, V.; Gouzy, J.; Young, N. D.; Prosperi, J.-M.; Leprince, O.; Thompson, R. D.; Buitink, J.; Burstin, J.; Gallardo, K. Genome-wide association studies with proteomics data reveal genes important for synthesis, transport and packaging of globulins in legume seeds. *New Phytol.* **2017**, *214*, 1597−1613.

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX

(32) Ohm, J.-B.; Manthey, F.; Elias, E. M. Variation and Correlation of Protein Molecular Weight Distribution and Semolina Quality Parameters for Durum Genotypes Grown in North Dakota. *Cereal Chem.* **2017**, *94*, 780−788.
(33) Ohm, J.-B.; Hareland, G.; Simsek, S.; Seabourn, B. Size-Exclusion HPLC of Protein Using a Narrow-Bore Column for Evaluation of Breadmaking Quality of Hard Spring Wheat Flours. *Cereal Chem.* **2009**, *86*, 463−469.
(34) Larroque, O. R.; Bekes, F. Rapid Size-Exclusion Chromatography Analysis of Molecular Size Distribution for Wheat Endosperm Protein. *Cereal Chem.* **2000**, *77*, 451−453.
(35) Oomah, B.; Voldeng, H.; Fregeau-Reid, J. A. Characterization of soybean proteins by HPLC. *Plant Foods Hum. Nutr.* **1994**, *45*, 251−263.
(36) Khazaei, H.; Stoddard, F. L.; Purves, R. W.; Vandenberg, A. A multi-parent faba bean (Vicia faba L.) population for future genomic studies. *Plant Genet. Resour.* **2018**, *16*, 419−423.
(37) Mertens, C.; Dehon, L.; Bourgeois, A.; Verhaeghe-Cartrysse, C.; Blecker, C. Agronomical factors influencing the legumin/vicilin ratio in pea (Pisum sativum L.) seeds. *J. Sci. Food Agric.* **2012**, *92*, 1591−1596.
(38) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248−254.
(39) Krishnan, H. B.; Oehrle, N. W.; Natarajan, S. S. A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: A case study using Glycine max. *Proteomics* **2009**, *9*, 3174−3188.
(40) Mosse, J. Nitrogen-to-protein conversion factor for ten cereals and six legumes or oilseeds. A reappraisal of its definition and determination. Variation according to species and to seed protein content. *J. Agric. Food Chem.* **1990**, *38*, 18−24.
(41) Le Signor, C.; Gallardo, K.; Prosperi, J. M.; Salon, C.; Quillien, L.; Thompson, R.; Duc, G. Genetic diversity for seed protein composition in Medicago truncatula. *Plant Genet. Resour.* **2005**, *3*, 59−71.
(42) Lampi, A.-M.; Yang, Z.; Mustonen, O.; Piironen, V. Potential of faba bean lipase and lipoxygenase to promote formation of volatile lipid oxidation products in food models. *Food Chem.* **2020**, *311*, 125982.
(43) Lee, K. J.; Hwang, J. E.; Velusamy, V.; Ha, B. K.; Kim, J. B.; Kim, S. H.; Ahn, J. W.; Kang, S. Y.; Kim, D. S. Selection and molecular characterization of a lipoxygenase-free soybean mutant line induced by gamma irradiation. *Theor. Appl. Genet.* **2014**, *127*, 2405−2413.
(44) Forster, C.; North, H.; Afzal, N.; Domoney, C.; Hornostaj, A.; Robinson, D. S.; Casey, R. Molecular analysis of a null mutant for pea (Pisum sativum L.) seed lipoxygenase-2. *Plant Mol. Biol.* **1999**, *39*, 1209−1220.
(45) Sęczyk, Ł.; Świeca, M.; Kapusta, I.; Gawlik-Dziki, U. Protein-Phenolic Interactions as a Factor Affecting the Physicochemical Properties of White Bean Proteins. *Molecules* **2019**, *24*, 408.
(46) Defaix, C.; Aymes, A.; Slabi, S. A.; Basselin, M.; Mathe, C.; Galet, O.; Kapel, R. A new size-exclusion chromatography method for fast rapeseed albumin and globulin quantification. *Food Chem.* **2019**, *287*, 151−159.
(47) O'Kane, F. E.; Happe, R. P.; Vereijken, J. M.; Gruppen, H.; van Boekel, M. A. J. S. Characterization of Pea Vicilin. 1. Denoting Convicilin as the α-Subunit of the Pisum Vicilin Family. *J. Agric. Food Chem.* **2004**, *52*, 3141−3148.
(48) Verni, M.; Coda, R.; Rizzello, C. G., Chapter 37- The Use of Faba Bean Flour to Improve the Nutritional and Functional Features of Cereal-Based Foods: Perspectives and Future Strategies. In *Flour and Breads and their Fortification in Health and Disease Prevention (Second Edition)*, Preedy, V. R.; Watson, R. R., Eds. Academic Press: 2019; pp 465−475.

https://dx.doi.org/10.1021/acs.jafc.0c02927
J. Agric. Food Chem. XXXX, XXX, XXX−XXX